

Relmecs, diciembre 2016, vol. 6, no. 2, e011, ISSN 1853-7863  
Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación.  
Centro Interdisciplinario de Metodología de las Ciencias Sociales.  
Red Latinoamericana de Metodología de las Ciencias Sociales

# Entre el método científico y la práctica del oficio. Estrategias de enseñanza de tabulación y análisis de datos en investigación social

Between the scientific method and the professional practice.  
Training strategies about data tabulation and analysis on social research

**Gonzalo Seid\***; **Patricia Fernández\*\***; **Jessica Malegarie\*\*\***

\*Facultad de Ciencias Sociales - Universidad de Buenos Aires. Becario del Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina | [gonzaloseid@gmail.com](mailto:gonzaloseid@gmail.com); \*\*Facultad de Ciencias Sociales, Universidad de Buenos Aires, Argentina | [pfernandez4@yahoo.com.ar](mailto:pfernandez4@yahoo.com.ar); \*\*\*Facultad de Ciencias Sociales, Universidad de Buenos Aires, Argentina | [jessi.malegarie@gmail.com](mailto:jessi.malegarie@gmail.com)

## PALABRAS CLAVE

Investigación social  
Estrategias de enseñanza  
Análisis cuantitativo

## RESUMEN

Este trabajo pretende reseñar algunas dificultades en el proceso de enseñanza-aprendizaje de análisis de datos cuantitativos en la investigación social. En relación con las distintas dificultades identificadas se consignan las estrategias didácticas empleadas para superarlas. Junto a la transmisión de los aspectos técnicos de los procedimientos de elaboración de cuadros y los recursos para el análisis –como las pruebas de hipótesis, los coeficientes de asociación y las lecturas porcentuales–, consideramos necesario el énfasis permanente en el papel de la teoría y los objetivos de investigación en las decisiones que se toman en el proceso investigativo.

## KEYWORDS

Social research  
Teaching strategies  
Quantitative analysis

## ABSTRACT

Some difficulties on teaching and learning the analysis of quantitative data on social research are expounded in this paper. We propose some didactical strategies used to overcome identified difficulties. Besides the teaching of technical considerations about tables' construction and tools for the analysis –such as hypothesis tests, association coefficients and percentage readings– we consider that it is necessary to emphasize the role of theory and research objectives on decisions taken on research process.

Recibido: 17 de agosto de 2015 | Aceptado: 9 de noviembre de 2015 | Publicado: 30 de diciembre de 2016

*Cita sugerida:* Seid, G.; Fernández, Patricia; Malegarie, Jessica. (2015). Entre el método científico y la práctica del oficio. Estrategias de enseñanza de tabulación y análisis de datos en investigación social. *Revista Latinoamericana de Metodología de las Ciencias Sociales*, 6(2), e011. Recuperado a partir de: <http://www.relmecs.fahce.unlp.edu.ar/article/view/relmecse011>



Esta obra está bajo licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional  
[http://creativecommons.org/licenses/by-nc-sa/4.0/deed.es\\_AR](http://creativecommons.org/licenses/by-nc-sa/4.0/deed.es_AR)

## 1. Introducción

En el proceso de investigación social la etapa del análisis constituye el momento en que se procura dar cumplimiento a los objetivos y responder a las preguntas-problema, para contribuir a la producción de conocimiento en determinada temática. Cuando se trabaja con un enfoque cuantitativo una de las herramientas de análisis más utilizadas es la lectura de cuadros bivariados y multivariados, a partir de los cuales se analiza el comportamiento de una población mediante la puesta en relación de dos o más variables. Las relaciones entre variables, pautadas a partir de hipótesis y plasmadas en tablas de contingencia, expresan vínculos entre conceptos del marco teórico. Desde un principio, la decisión de poner en relación determinadas variables –y no otras– se sustenta en expectativas teóricas.

La metodología, conformada por procedimientos y métodos para la construcción de la evidencia empírica, se basa en paradigmas epistemológicos y reflexiona sobre el papel de la teoría en relación a la evidencia que se produce (Sautu, 2005 y 2010). Por esta razón, consideramos de fundamental importancia en la enseñanza del análisis de datos cuantitativos transmitir una serie de contenidos en torno al anclaje de las técnicas en el proceso teórico-metodológico más amplio que le da sentido. Es decir, reafirmar en la práctica que toda construcción de un cuadro y su posterior interpretación implican decisiones que son tanto teóricas como metodológicas. La explicitación permanente del papel de la teoría, del problema y de los objetivos de investigación constituye una tarea fundamental frente a las dificultades encontradas en el aprendizaje del análisis de datos.

Considerando los aspectos mencionados, el propósito de este artículo es presentar un conjunto de reflexiones sobre el proceso de enseñanza-aprendizaje del análisis de datos dentro de un abordaje cuantitativo de investigación. Dichas reflexiones se encuadran en nuestra experiencia como docentes de la materia Metodología de la Investigación Social de la carrera de Sociología de la Universidad de Buenos Aires. Nos proponemos reflexionar en torno a las dificultades que se manifiestan en el proceso de aprendizaje de la construcción y análisis de tablas bi y multivariadas, así como sobre las estrategias de enseñanza desarrolladas para abordar dichas problemáticas. En un primer apartado, se exponen los señalamientos introductorios que se transmiten a los estudiantes explicitando para qué se realiza y en qué consiste el análisis de tablas en el marco de un proceso de investigación social. En un segundo apartado, se desarrollan los aspectos relativos a los procedimientos y recursos para el análisis –prueba de hipótesis, coeficientes de asociación, lectura porcentual– haciendo foco también en las dificultades de aprendizaje y en las estrategias de enseñanza utilizadas para lograr transmitir la primacía de las decisiones teórico-metodológicas sobre las consideraciones técnicas. En la última sección, se trabajan los desafíos y las complejidades particulares de la enseñanza-aprendizaje de tablas multivariadas.

## 2. Algunas nociones preliminares para abordar la construcción y análisis de cuadros

### *Hipótesis de trabajo*

Una primera dificultad que aparece al momento de enseñar a realizar un análisis adecuado de tablas de contingencia es que los estudiantes no siempre tienen en claro el objetivo de su construcción. En algunos casos seleccionan variables sin articular ni contextualizar la lectura dentro del proceso de investigación que se ha emprendido y sin basar la construcción de la tabla en un supuesto teórico específico. Dicho en otros términos, cuando se decide observar la relación entre dos o más variables los estudiantes muchas veces no logran considerar la relevancia teórica de dicha relación y la necesaria existencia de alguna hipótesis de trabajo que la sustente.

Un primer recurso pedagógico en estos casos consiste en trabajar sobre las distintas hipótesis que podrían

establecerse y resaltar que son estas las que darán marco y sostén a lo que se desea analizar, para evitar construir tablas que carezcan de sentido teórico. Basar el análisis de las relaciones entre variables en una hipótesis es uno de los elementos que marcan la diferencia entre un acercamiento desde el sentido común y un análisis sociológico. Asimismo, se transmite que estas hipótesis no necesariamente deben ser planteadas al inicio del proceso en tanto respuestas anticipadas a las preguntas-problema, sino que pueden ser hipótesis más puntuales que se formulan en el transcurso de la investigación, en vinculación directa con las tablas que se elaboran.

Otra dificultad habitual entre los estudiantes es que, aún analizando la relación entre variables a partir de un encuadre teórico, suelen limitar el análisis a un determinado procesamiento de las variables inicialmente planteado. En este sentido, un aporte didáctico importante es el de remarcar que pueden construirse distintos sistemas de categorías para una variable, según decisiones teórico-metodológicas. Se señalan las distintas posibilidades de agrupar y reagrupar las categorías de las variables. Esto implica la posibilidad de realizar distintos procesamientos y compararlos para lograr mayor riqueza analítica. Las variaciones en las distribuciones de los casos en los espacios de propiedades generados suponen que las hipótesis pueden corroborarse en distintos grados cuando se trabaja con unos u otros sistemas de categorías. Las posibilidades que se abren son múltiples y pueden reverse o ajustarse varias veces hasta encontrar la formulación más satisfactoria.

### *Plan de análisis*

Para definir qué variables y categorías han de ser puestas en relación en las hipótesis –y por lo tanto en las tablas de contingencia– se propone trabajar con los estudiantes detenidamente en la construcción de un plan de análisis. El plan de análisis cuantitativo contiene un esquema sobre el conjunto de cuadros esperados, es decir, las decisiones relativas a todas las tablas que se confeccionarán, formalizando las hipótesis en torno a las cuales han de ser interpretados los datos producidos. Su construcción implica una labor de previsión sobre las estrategias de análisis y el detalle de procedimientos que deben llevarse a cabo para cumplir los objetivos. Además, es posible anticipar cómo se espera que se ubiquen las unidades de análisis en los espacios de propiedades definidos por las hipótesis (Barton, 1973; Mora y Araujo, 1984). Cabe aclarar que estas decisiones están condicionadas por las variables que hayan sido incluidas en el instrumento de recolección de información y por la existencia de una suficiente cantidad de casos que permita analizar su distribución en los espacios de propiedades generados.

En el proceso de enseñanza se procura enfatizar que cada tabla que se elabora sólo tiene sentido cuando existe una justificación teórica acerca de lo que espera observarse. Si en lugar de dar formato de cuadro a las hipótesis, se procediera observando todos los cruces posibles de variables, ésta sería una tarea no sólo innecesaria y engorrosa, sino ante todo carente de relevancia teórica. Tendría como resultado datos desconectados de la teoría, que por sí solos no pueden constituir un avance en la producción de conocimiento científico. Los programas informáticos para el tratamiento estadístico de bases de datos permiten sin grandes dificultades “cruzar todo con todo” y eventualmente podrían servir para explorar relaciones impensadas en los datos, pero se advierte que de ninguna manera se estaría produciendo conocimiento científico si se realizara una lectura de datos desprovista de la teoría que le da sustento. Como contrapartida, se hace notar que tampoco es adecuado forzar la lectura teórica para que los datos confirmen lo que “queremos que digan”. La confrontación entre teoría y empiria es lo que garantiza un adecuado proceso de investigación. Al momento de enseñar la etapa de análisis, la construcción previa del plan de cuadros como expresión de hipótesis resulta necesaria para poner de manifiesto que existen decisiones teóricas tras su construcción.

El plan de cuadros también es útil para que los estudiantes logren comprender que la contrastación de cada hipótesis permite avanzar en la producción de conocimiento, pero no agota en sí mismo los objetivos de investigación, sino que es el conjunto de las hipótesis de trabajo contrastadas en las tablas que se producen, con la articulación teórica y la ilación lógica que le confieren quienes investigan, lo que puede dar respuesta a los interrogantes. La lectura de un cuadro no debe ser entendida entonces como un fin en sí mismo sino como un medio para el análisis e interpretación de los resultados. No debe perderse de vista el panorama general del plan de análisis, de modo que los avances parciales en el conocimiento a partir de la lectura de cada cuadro puedan ser articulados de un modo coherente y enfocado en dar cumplimiento a los objetivos.

### *Revisión de la construcción de las variables*

Cuando se analizan los resultados, en ocasiones la denominación de las variables puede hacer perder de vista cómo han sido construidas a partir de la información relevada, por ejemplo, cómo se preguntó en un cuestionario. Si bien suele resultar innecesario reiterar las limitaciones de una fuente de datos cada vez que se analiza un cuadro, puede ser provechoso explicitarlas toda vez que su lectura produce la impresión de que si se hubieran construido las variables de otra manera o con otra técnica los datos apoyarían interpretaciones muy distintas. La reflexión debe apuntar a dilucidar si esa impresión se debe a un prejuicio porque los datos no se ajustan a las expectativas o si debe problematizarse la construcción de esos datos. Puesto que los datos son construcciones conceptuales creadas en el marco de una investigación particular, con objetivos y fines específicos, es necesario mantener un ejercicio constante de vigilancia epistemológica, de cara a repensar su pertinencia en función de cada contexto de investigación (Bourdieu, Passeron y Chamboredon, 2011).

### *Análisis cuantitativo y recursos estadísticos*

Varias de las dificultades en el aprendizaje del análisis cuantitativo se vinculan a la confusión de éste con las herramientas estadísticas que se utilizan. En no pocas ocasiones los recursos estadísticos son erróneamente considerados como si hicieran posible conducir la investigación mediante meros procedimientos técnicos e impersonales. Para evitar esta concepción de la técnica, el rol del docente se centra en distinguir análisis cuantitativo de análisis estadístico, teniendo presente que aunque los datos puedan requerir un tratamiento estadístico, éste debe subordinarse al abordaje teórico-metodológico. El investigador no puede tampoco llevar a cabo su trabajo sin conocimiento de estadística y delegar en un especialista el análisis estadístico, sino que los problemas del análisis han de anticiparse, influyendo las consideraciones estadísticas a lo largo de todo el proceso de investigación (Blalock, 1986: 19). En definitiva, dentro del proceso de enseñanza-aprendizaje se remarca fuertemente que el investigador se vale de los recursos estadísticos pero no puede dejar de tomar decisiones teórico-metodológicas indelegables si quiere producir conocimiento relevante para llevar a buen puerto la investigación. En caso de que no quede lugar para tales decisiones, el riesgo es caer en lo que Bourdieu, Passeron y Chamboredon (2011) denominaban la autonomización del método, que ocurre cuando el desarrollo de técnicas estadísticas y modelos matemáticos se convierten en el objeto mismo de las investigaciones, y los procesos sociales quedan reducidos a relaciones estadísticas entre variables. Teoría, metodología y técnicas de análisis deben articularse reflexivamente, atendiendo a la coherencia y ejerciendo vigilancia epistemológica (Scribano, 2008).

## **3. Recursos y procedimientos para el análisis de tablas**

### *Implicancias de la significación estadística*

Luego de haber elaborado las hipótesis a partir de la teoría y haber efectuado el procesamiento del que resulta una distribución de casos en los espacios de propiedades, es el momento para el análisis de la tabla. Es habitual comenzar esta tarea realizando una prueba de hipótesis. Cuando se utilizan estas pruebas es necesario evaluar su alcance a los fines de la investigación, para evitar una aplicación mecánica y metodológicamente irreflexiva. En el momento de enseñanza del uso de una prueba de hipótesis como chi cuadrado, se procuran transmitir las consideraciones técnicas necesarias que sirven para definir el escenario estadístico en que se trabaja. En este punto nuevamente se refuerza la idea de que el uso de técnicas estadísticas sólo es un medio para realizar el análisis desde un abordaje teórico-metodológico, puesto que la relevancia teórica no depende del resultado del test estadístico.

Una de las principales dificultades en su aprendizaje reside en entender qué implica que una hipótesis no sea estadísticamente significativa para determinado nivel de confianza, por ejemplo, si esto supone que necesariamente deba aceptarse la hipótesis nula o que haya que dar por finalizado el análisis de la relación. La ausencia de significación estadística puede deberse a varios motivos, tales como el modo de construir las variables, una insuficiente cantidad de casos, los grados de libertad del cuadro, las características de la estrategia de medición, etc. Por ello, se intenta transmitir que la ausencia de significación estadística entre las variables analizadas puede ser tan relevante como la determinación de su existencia. A los fines sociológicos, la situación en que una prueba como el chi cuadrado no asume significación estadística no menoscaba la relevancia teórica de una hipótesis y, en contrapartida, el hecho de que una relación entre variables sea estadísticamente significativa no es suficiente por sí mismo para producir conocimiento científico.

### *Interpretación de los coeficientes de asociación*

El uso de pruebas de hipótesis para determinar la existencia de asociación estadísticamente significativa suele complementarse con el análisis de los coeficientes de asociación, que permiten determinar la intensidad de la relación entre variables. Nuevamente, al enseñar cómo seleccionar e interpretar adecuadamente un coeficiente, se procura lograr no sólo el dominio técnico sino también un uso reflexivo desde los objetivos de una investigación.

En ocasiones, las dificultades aparecen respecto de la forma rinconal o diagonal de las hipótesis, como consideración necesaria para la elección del tipo de coeficiente a utilizar. Asimismo, no siempre resulta claro que el tipo de hipótesis planteada es independiente de lo observado en el cuadro. Frente a estas dificultades una estrategia muy útil consiste en trabajar reiteradamente sobre diferentes formulaciones de hipótesis (rinconales, diagonales, irregulares) asociadas a una misma tabla de contingencia, lo que permite observar que pueden corroborarse o refutarse distintas hipótesis frente a una misma distribución porcentual observada en un cuadro bivariado.

Un aspecto central a destacar en la enseñanza de coeficientes de asociación se relaciona con la interpretación de su valor resultante. Una vez que los estudiantes han aprendido que a medida que el valor del coeficiente se aleja de cero está expresando cierta intensidad de relación entre las variables, deben ser capaces de evaluar dicha intensidad en cada situación. Con respecto a las consideraciones técnicas, se señala que en la interpretación de los valores de coeficientes curvilineales, como Q de Yule o Gamma, deben tener presente que tienden a asumir valores más altos puesto que los mismos se basan en una definición menos restrictiva de asociación perfecta. La misma refiere a una hipótesis rinconal que puede cumplirse a la perfección poniendo una sola condición: para que haya asociación perfecta en una tabla de dos por dos bastaría con que un solo espacio tenga cero frecuencias. Inversamente, los coeficientes lineales, así como la V de Cramer,

tienden a “subestimar” el grado de relación por suponer una definición exigente de asociación perfecta (García Ferrando, 1985).

Más allá de conocer los cálculos que supone cada coeficiente, otro tipo de cuestiones deben atenderse para una adecuada interpretación de los mismos. En no pocas ocasiones los estudiantes tienden a hacer un uso mecánico de los coeficientes y una lectura demasiado simple de una relación bivariada como si fuera en sí misma “fuerte o débil”. La interpretación individual de cada resultado en una escala de mayor o menor fuerza de asociación reduce la comprensión global de un fenómeno a la luz del impacto de diferentes variables y situaciones. Por ello, un recaudo que se intenta transmitir consiste en que el valor no sea interpretado aisladamente como alto, medio, bajo, etc., sino que su interpretación esté fundada en la realización de comparaciones, para las cuales se puede tomar como referencia el valor del mismo coeficiente en otros casos o para otras hipótesis.

Consideramos recomendable interpretar el valor de un coeficiente a la luz del valor que el mismo adquiere para la relación de la misma variable dependiente con distintas variables independientes. Para que las comparaciones respecto de la fuerza de incidencia de distintas variables independientes sean correctas es necesario realizarlas con el mismo coeficiente, utilizando así el mismo criterio de medida en todos los casos. Por otro lado, el conocimiento teórico acumulado sobre el comportamiento del fenómeno también contribuirá a determinar cuán alto o bajo es el resultado, en relación con expectativas teóricas o derivadas de la experiencia previa. En cualquier caso, la comparación es un paso nodal en el análisis, que posibilita que los recursos técnicos sean subordinados a las decisiones teórico-metodológicas, contribuyendo a la producción de conocimiento teóricamente relevante.

### *La lectura porcentual como recurso central del análisis*

Las técnicas anteriormente mencionadas –prueba de hipótesis y coeficientes de asociación– son presentadas en la enseñanza como complementos de la lectura porcentual de las tablas. El análisis de porcentajes es un recurso central que posibilita una comprensión más profunda de la relación entre las variables analizadas. Mientras que las pruebas de hipótesis y los coeficientes permiten un conocimiento sintético acerca de la relación entre variables, la lectura porcentual de tablas de contingencia permite un análisis cuantitativo más refinado y analítico del mismo.

Aunque un programa estadístico pueda elaborar las tablas que se le ordenen y calcular los porcentajes, resulta útil para la enseñanza ejercitar el pasaje de la matriz de datos a los cuadros de frecuencias absolutas y con éstos, las distintas posibilidades de porcentualización. Si bien el cálculo de porcentajes es una operación sencilla para los estudiantes, resulta fundamental que se comprenda cabalmente su utilidad para realizar comparaciones válidas con mayor facilidad, en tanto constituyen una manera de estandarizar las frecuencias condicionales para liberarlas de los efectos de las diferencias de los marginales, haciendo comparable una parte respecto al todo (Baranger, 1999).

Asimismo, es necesario tener en cuenta sobre qué total de casos han sido calculados los porcentajes, ya sea para poder re-porcentualizar agrupando categorías o, más importante, para tener los debidos recaudos en la interpretación de porcentajes calculados sobre pocos casos. La cantidad de decimales con la cual puede trabajarse depende de la cantidad de casos, lo cual debe tenerse en cuenta para evitar que los porcentajes aparenten un nivel de exactitud que no pueden tener cuando la muestra es pequeña. Para facilitar el análisis se sugiere a los estudiantes presentar los porcentajes sin decimales o con un solo decimal.

Una de las dificultades que surgen en la elaboración de las tablas de contingencia se refiere al sentido en que han de calcularse los porcentajes según los roles de las variables. Para ello, creemos que resulta

imprescindible la comprensión de por qué los porcentajes deben calcularse en un sentido y no en otro si de lo que se trata es de corroborar o refutar hipótesis. Un recurso para lograr dicha comprensión consiste en que los estudiantes reflexionen sobre el significado que tienen los porcentajes, mediante la prueba y comparación de valores porcentuales calculados sobre distintos totales o subtotales. De este modo, mediante sucesivas ejercitaciones, se espera que los estudiantes lleguen a comprender que:

-Si se procura conocer qué porcentaje de la población puede ser clasificado en varios atributos simultáneamente –por ejemplo, varón y de nivel educativo alto– es necesario calcular los porcentajes sobre el total de la tabla. Esto expresa cómo se distribuyen los espacios de propiedades sobre el total, lo que puede utilizarse para describir el porcentaje de la muestra que posee esa doble condición o comparar con otra situación similar, otro momento, etc.

-Si se busca describir perfiles de la población y subpoblaciones, puede ser útil utilizar los porcentajes “de filas”, es decir, aquellos calculados en el sentido de la variable dependiente. Estos porcentajes sólo pueden servir a los fines descriptivos univariados de las características de distintos subgrupos. Se trata de examinar la distribución de la variable de estudio en función de otra/s variable/s pero sin asumir una relación de dependencia –por ejemplo, qué porcentaje de los menores de 35 años que tienen nivel educativo alto son varones–, brindando así una idea de la composición interna o perfil de cierto subgrupo.

-Para realizar una lectura en función de una hipótesis bivariada es necesario calcular los porcentajes en el sentido de la variable independiente y leerlos comparando en el sentido de la variable dependiente (Zeisel, 1974). De este modo, se puede analizar la influencia de la variable independiente sobre la dependiente. Por ejemplo, saber qué porcentaje de los varones tiene nivel educativo alto, al comparar con el porcentaje que tienen las mujeres en dicho nivel educativo, permite a partir de las diferencias porcentuales aproximarse a la incidencia de la variable sexo en el nivel educativo alcanzado. Esta distribución porcentual bivariable pone en relación las distribuciones condicionales con una distribución global de la variable dependiente. La distribución de la variable dependiente dentro de las categorías de la variable independiente permite examinar las condiciones que se supone que influyen en el modo de dicha distribución.

Para facilitar su comprensión, en la enseñanza se transmite que la variable independiente puede ser entendida como condición que hace probable cierto resultado, por lo cual es necesario comparar diferentes condiciones –categorías de la variable independiente– para resultados equivalentes –de una misma categoría de la variable dependiente–, examinando qué le ocurre a la variable de estudio cuando se le “aplica” cada una de las categorías de la variable independiente.

Cuando se trata de una tabla de distribución conjunta de dos variables dicotómicas –usualmente denominados cuadros de dos por dos– bastará con comparar dos porcentajes para leer el cuadro por completo, puesto que habiendo leído los dos porcentajes de una categoría de la variable dependiente, los de la otra categoría de la variable dependiente son complementarios. La diferencia entre los dos porcentajes comparados será la diferencia porcentual o *épsilon*, que en este caso sintetiza todo el cuadro, indicando la fuerza de la relación entre ambas variables (García Ferrando, 1985).

En los cuadros de más de dos por dos, ya no habrá una única diferencia porcentual sino varias y por ende ninguna de ellas sintetizará en una única medida la fuerza la relación. Cuantas más sean las categorías de las variables, mayores son las posibilidades de comparaciones que se abren y por ende, son mayores las dificultades que se les presentan a los estudiantes. Se señala entonces que probablemente no sea necesario leer todos los porcentajes cuando haya muchas categorías, ya que serán siempre las hipótesis las que orientarán la lectura indicando qué espacios de propiedades son los más relevantes según las expectativas teóricas –lo que no significa que la lectura deba restringirse solamente a las celdas verificadoras de la hipótesis–.



Se destaca además que todo tipo de situaciones intermedias, variaciones sutiles y matices respecto del modo de relación entre las variables pueden ser descriptos pormenorizadamente mediante la lectura porcentual, lo que permite analizar en profundidad la información resultante de un cuadro, algo que no es posible realizar sólo con los coeficientes de asociación.

Las dificultades de los estudiantes en lo que respecta a la lectura y análisis de cuadros abarcan desde el sentido en el cual calculan los porcentajes hasta la realización de lecturas erróneas con respecto a dichas porcentualizaciones. Por este motivo, resulta óptimo para la asimilación del tema, dedicar el mayor tiempo que el currículo permita a realizar distintos tipos de ejercitaciones. Las mismas pueden abarcar instancias individuales, conjuntas, verbales, por escrito y puestas en común, para superar las distintas limitaciones que surgen en el aprendizaje: análisis realizado correctamente en forma verbal pero con dificultad de plasmarlo por escrito, pormenorizado análisis escrito que pierde el sentido orientador de la lectura basada en la hipótesis, lecturas incorrectas que llevan a conclusiones equivocadas, por poner algunos ejemplos. Lo que se busca en definitiva es la apropiación real del uso de una técnica que será seguramente una de las que con más frecuencia utilicen los estudiantes a lo largo de la carrera y de su desempeño profesional.

#### *Interpretaciones y conjeturas más allá de los datos*

Antes de pasar al apartado siguiente, nos parece importante destacar un aspecto adicional del aprendizaje del análisis de tablas que, más allá de las cuestiones técnicas, vinculadas a aprender a percibir y dotar de sentido lo que “dice” el cuadro, implica a toda la información que el cuadro tiene para proporcionar. Seguramente algunas sutilezas sólo podrían resultar sugerentes para quienes cuentan con mayor experiencia realizando investigaciones, pero vale la pena tener en cuenta que muchas veces los aspectos más interesantes del proceso de producción de conocimiento pueden surgir de las maneras menos esperadas. Por este motivo, aunque las tablas se elaboren habitualmente para verificar hipótesis, también debe considerarse la posibilidad de que el cuadro pueda ser utilizado como elemento sugerente de nuevas hipótesis “más allá de lo que dice el cuadro”, con procesos o variables no relevadas que contribuyen a la interpretación de los datos. Una distribución o alguna celda que se aleja de lo esperado, acompañada de una mirada aguda, suele disparar hipótesis de por qué ocurre esto, conjeturas que suelen emerger casi espontáneamente del “sentido común sociológico” o del conocimiento teórico acumulado. Pueden resultar particularmente fructíferas las ideas que emergen en las discusiones grupales, ya que pueden ser tomadas como fuente de hipótesis o bien cuestionadas por basarse en preconcepciones que requieren ser explicitadas y sobre las que debe reflexionarse con actitud de vigilancia epistemológica.

#### **4. Análisis de relaciones multivariadas**

En este apartado, desarrollaremos una serie de aspectos que consideramos relevantes en la enseñanza del análisis multivariado mediante tablas de contingencia. Las tablas constituyen la aproximación más básica y tradicional al análisis multivariado, y por eso la más apropiada para un curso introductorio de estudiantes de grado. “El análisis multivariable implica introducir terceras y cuartas variables en la tabla, habitualmente no más, pues el cruce entre ellas aumenta el número de casillas con la consecuente insuficiencia de frecuencias en ellas y la complejidad de la lectura de la información” (López Roldán y Fachelli, 2015). Antes de introducirnos en este tema, cabe destacar que también se menciona en las clases la existencia de otras herramientas para analizar simultáneamente un conjunto de variables nominales, tales como los modelos log-lineales o el análisis de correspondencias múltiples.

Se pueden utilizar modelos log-lineales para analizar simultáneamente las asociaciones e interacciones entre



tres o más variables categóricas. En estos modelos todas las variables involucradas actúan como variables independientes y la variable dependiente es el número de casos en una de las casillas de la tabulación cruzada. Se analizan todas las interacciones posibles entre las variables, observando la existencia de relaciones significativas entre las mismas (Federico y Oliva, 1997). La finalidad es seleccionar el modelo que incluya menos efectos –el más parsimonioso– y cuya tabla de frecuencias esperadas reproduzca las frecuencias observadas –el mejor ajuste– (Latiesa, 1991).

Otra posibilidad de uso muy difundido para el análisis multivariado de variables nominales es el Análisis de Correspondencias Múltiples, que permite encontrar factores que expresan combinaciones de las variables originales. Esta técnica reduce la información, pero permite ganar en significado al identificar los principales factores de diferenciación, ordenados jerárquicamente según la inercia explicada. Asimismo, la representación gráfica contribuye a la interpretación de las correspondencias o asociaciones según la proximidad espacial de las categorías de las variables. Como la técnica trabaja a partir de las categorías, debe tenerse en cuenta que la inercia explicada por una variable es mayor cuantas más categorías tenga. El análisis de correspondencias ha sido utilizado en *La Distinción* por Pierre Bourdieu, quien consideraba que los supuestos de esta técnica se adecuaban a su perspectiva teórica para observar la correspondencia u homología entre la estructura de posiciones de clase y la estructura de estilos de vida. Se trata de una técnica descriptiva que en lugar de segmentar el objeto en variables, toma todos los factores simultáneamente, de modo tal que el peso de cada relación es magnetizado por el de todas las demás relaciones. Esto permitiría representar en el espacio factorial la complejidad de un espacio social multidimensional (Baranger, 2004).

### *Salto en complejidad*

Una vez que se ha corroborado que existe relación entre dos variables, es habitual la decisión de introducir otras variables de control para examinar si esa asociación inicial persiste, se modifica o desaparece. Cuando los estudiantes dominan los conocimientos básicos sobre la construcción y lectura de cuadros bivariados, la incorporación de terceras variables genera otro nivel de complejidad y esto se vislumbra en varios sentidos: el desafío de pensar simultáneamente el inter-juego de las variables, los distintos roles que asumen las variables de acuerdo a su orden temporal, el sentido en que se calculan los porcentajes, la creación de nuevos espacios de propiedades y frecuencias porcentuales que han de ser interpretadas, y desde luego la relevancia teórica que deberá fundamentar la elección de cada nueva variable que se incorpore al análisis.

El cuadro de tres o más variables –particularmente cuando estas tienen más de tres categorías– genera en ocasiones desorientación en los estudiantes, incluso cuando ya dominan los aspectos procedimentales de la lectura de cuadros. Suele ocurrir que en las primeras ejercitaciones se lean los cuadros de la misma manera que en tablas bivariadas, por lo cual no se logra afirmar nada sobre qué cambia o se mantiene ante la introducción de la variable de control. La multiplicidad de posibilidades de comparaciones hace que no sea tan sencillo aprender a decidir qué comparar con qué y cómo darse cuenta si las diferencias se deben al efecto de la variable de control o a cambios en las frecuencias marginales. La mayor complejidad también se debe a que no sólo es necesario efectuar comparaciones entre porcentajes, sino también comparar las propias comparaciones al examinar si las diferencias porcentuales varían ante las nuevas condiciones definidas por las categorías de la variable introducida. Asimismo, es importante considerar frente al cruce de tres o más variables con varias categorías, la cantidad de casos absolutos que queda en cada casilla, de modo de hacer lecturas sobre un número suficiente de unidades de análisis, para evitar sesgos en la subpoblación.

### *Lógica experimental en el análisis multivariado*

Si bien el análisis multivariado no sólo se utiliza en diseños explicativos, sino también cuando los objetivos son descriptivos o incluso exploratorios (Lago Martínez, Mauro y Álvarez, 2000), la lógica explicativa siempre requiere de un análisis multivariado que controle las relaciones postuladas en las hipótesis originales. Por tratarse de un recurso de análisis que ha sido pensado como alternativa al diseño experimental clásico, la presentación conjunta de la lógica experimental y el análisis de tablas multivariadas se emplea como estrategia pedagógica (Cohen y Gómez Rojas, 2003).

La enseñanza de la lógica experimental, en relación con los desafíos para emplearla en ciencias sociales, constituye uno de los primeros pasos necesarios para el aprendizaje de la temática. En esta etapa se señalan las diferencias entre el experimento clásico y las aproximaciones que se realizan desde las ciencias sociales (Hyman, 1984). En ambos casos existen al menos dos grupos de comparación, se manipulan variables y el orden temporal asume un papel importante en el análisis. La diferencia fundamental entre uno y otro modelo radica en que en ciencias sociales generalmente el control se efectúa *ex post facto*, es decir equiparando atributos luego de acaecidos los sucesos, y más precisamente, en el momento del análisis de la información relevada (Campbell y Stanley, 1973: 134). Así, el análisis multivariado dentro de los denominados "cuasi-experimentos" se refiere a diseños de investigación en los cuales el control no se efectúa por la selección de los sujetos o grupos de tratamiento y control. En estos diseños el grupo de tratamiento sirve como su propio control, al mantener constante el efecto de las variables de control.

### *Control de variables*

En relación con el aprendizaje de la lógica experimental, se desarrolla el concepto de *control de variables*, uno de los que presenta mayor complejidad para los estudiantes. Se requiere la comprensión cabal sobre qué es una variable de control, qué implica y qué hipótesis la respalda, para un adecuado análisis ulterior de sus efectos sobre la relación original. Tanto en la hipótesis original que genera la relación bivariada, como en cada una de las nuevas hipótesis que buscarán una explicación alternativa al fenómeno, la teoría estará presente. Por ello, el primer paso en el proceso de enseñanza consiste en exponer el nuevo desafío vinculado a pensar simultáneamente el inter-juego de variables y explicitar que introducir una variable de control en el análisis de la relación implica mantenerla constante. En este proceso de neutralizar los efectos de la tercera variable, se visualiza justamente si –y de qué forma– la variable de control genera modificaciones en la relación original.

Con la introducción de sucesivas variables de control y la consecuente identificación de las condiciones bajo las cuales tiene lugar un fenómeno, es posible aproximarse a una explicación. Si las hipótesis han sido adecuadamente planteadas a partir de la teoría, es probable que ya se haya conjeturado algún mecanismo que vincule las variables, el cual puede estar explicitado y clarificado en mayor o menor medida. En cualquier caso es necesaria la reflexión teórica permanente acerca de cómo se vinculan las variables. Las dificultades que suelen surgir entre los estudiantes se refieren precisamente a la falta de sustento teórico en la selección de las variables para la construcción de hipótesis alternativas. Nuevamente, se refuerza la idea de que la elección de la variable de control es una decisión teórica que desprende una hipótesis, ahora de carácter multivariado.

Se señala también la importancia de explicitar el rol que asumirá cada una de las variables en la relación (dependiente, independiente, de control antecedente o interviniente) y cuál es el orden temporal con respecto a las demás variables. Sobre estos dos aspectos se trabaja constantemente mediante distintos ejemplos, en

pos de reafirmar que constituyen decisiones que no son simplemente operativas sino fundamentalmente teórico-metodológicas.

### *Porcentualización según los roles de las variables*

Otro de los aspectos a destacar en la enseñanza de la temática es que a menudo los estudiantes no saben cómo proceder durante el análisis una vez que han elegido la variable a incorporar y han construido la hipótesis alternativa. En algunos casos, confunden el lugar donde debe estar ubicada la tercera variable y por lo tanto se les dificulta la construcción de una tabla acorde al rol que ocupa la variable introducida en la relación original.

Una vez comprendida la regla de calcular los porcentajes en el sentido de la variable independiente para examinar lo que ocurre con la hipótesis en tablas bivariadas, es necesario pasar al aprendizaje de cómo calcular los porcentajes ante la introducción de la variable de control. Puesto que un error común en el aprendizaje es calcular porcentajes como si la variable de control fuera una nueva variable independiente, resulta de utilidad enfatizar que la tercera variable se introduce en el análisis para examinar qué ocurre con la hipótesis original cuando se mantienen constantes los valores de esta tercera variable. Por este motivo es que la tabla bivariada debe replicarse para cada una de las categorías o valores de la variable de control (García Ferrando, 1985). Un aspecto al que conviene prestar atención es la ambigüedad de la habitual expresión “introducir” para referirse a la variable de control, puesto que la misma puede dificultar la comprensión de que el propósito de su introducción en el análisis es precisamente el de neutralizar o “quitar” sus posibles efectos. En este sentido, se señala que la introducción de una variable de control pone el foco precisamente en la relación bivariada original, pues se trata de analizarla desde la “lupa” de cada una de las categorías de la tercera variable.

### *Reducir la complejidad del cuadro*

Para analizar los datos que se obtuvieron en la tabulación cruzada bivariada con los que se obtengan al agregar sistemáticamente terceras variables, se comparan los coeficientes y las distintas lecturas porcentuales y se evalúa si la relación original fue afectada o no por la introducción de una nueva variable. Con la introducción de la variable de control en la tabla, la cantidad de espacios de propiedades se duplican, triplican, etc. acorde a la cantidad de categorías de esta tercera variable. Las múltiples posibilidades de comparaciones a realizar hacen que no sea tan sencillo aprender a decidir qué comparar y cómo darse cuenta si las diferencias se deben al efecto de la variable de control. Para que la lectura de tablas multivariadas sea lo más fructífera posible, resulta primordial ser capaz de reducir adecuadamente la complejidad resultante, cuidando a la vez de no desaprovechar su riqueza analítica por simplificar en exceso. Esto implica aprender a seleccionar los porcentajes más relevantes para ser leídos.

Una estrategia para ayudar a la comprensión en las primeras etapas del aprendizaje puede consistir en reducir la cantidad de categorías de cada cuadro, trabajando en un principio con cuadros de dos o tres categorías, de modo que las diferencias porcentuales en cada relación parcial puedan ser fácilmente comparadas entre sí. De todos modos, es también necesario que los estudiantes puedan aprender a lidiar con la complejidad de analizar cuadros más grandes, teniendo siempre presente que la decisión teórico-metodológica relativa al sistema de categorías de las variables no puede subordinarse al pragmatismo que implica hacer un cuadro más manejable e intuitivo para el análisis. Es importante señalar que no se trata de poner el énfasis en la cantidad de espacios de propiedades que se generan, sino en detectar cuál es la relación que existe por detrás del cruce entre las tres categorías de variables.

Principalmente cuando se trata de cuadros grandes, resulta de máxima utilidad antes de analizar las frecuencias condicionales, proceder a la lectura de los marginales de las tablas para cada categoría de la variable de control y utilizar alguna medida resumen o coeficiente que sintetice y permita comparar prontamente el grado de asociación en cada relación parcial. Así, las diferencias en las distribuciones entre la relación bivariada y las relaciones parciales para cada valor de la variable de control, guardarán correspondencia con la comparación de los coeficientes que se utilicen. Por ejemplo, si la introducción de una variable de control no genera ningún efecto sobre la relación bivariada, notaremos también que la comparación de los coeficientes no mostrará grandes diferencias. Por su parte, si la introducción de la variable de control modifica sustancialmente la distribución porcentual de la relación bivariada, notaremos amplia diferencia en el valor del coeficiente.

### *Selección de porcentajes relevantes para ser comparados*

En los cuadros con variables de más de tres categorías, no suele ser necesaria ni recomendable la lectura de cada uno de los porcentajes de los espacios de propiedades, pues se estaría repitiendo lo que ya dice el cuadro. En esta situación un interrogante común entre los estudiantes es ¿cómo saber cuáles merecen leerse o destacarse y cuáles pueden omitirse en un informe? La respuesta desde la experiencia docente, si bien varía en cada caso concreto, tiende nuevamente a enfocarse en la hipótesis. La misma permite orientar la lectura hacia aquellos espacios de propiedades que se consideren más relevantes para determinar si tiende a corroborarse o no. Asimismo, la distribución de los casos en la tabla bivariada original también ayuda a “intuir” qué celdas de las relaciones parciales pueden ser más relevantes para comparar, por ejemplo, aquellas con altas o bajas concentraciones de casos que requieren examinar si las mismas se mantienen para las distintas condiciones de la variable de control.

### *Situaciones posibles al introducir la variable de control en el análisis*

Para que los estudiantes puedan aprender a arribar a alguna conclusión, se transmiten las posibilidades más conocidas respecto de lo que puede ocurrir con la introducción de la variable de control, aclarando siempre que se trata más bien de situaciones “extremas” que no son las más frecuentes en ciencias sociales. Una primera posibilidad consiste en que la introducción de la nueva variable no modifique notablemente la relación bivariada y se sostenga la hipótesis original, de manera que resulte ajena la variable de control. Otra situación es que la relación bivariada se mantenga o fortalezca para una o algunas categorías de la variable de control y no para otras, es decir, que la hipótesis se especifique (se ven modificaciones en los coeficientes y porcentajes entre categorías de la variable de control). Una tercera posibilidad es que la introducción de la nueva variable en el análisis disuelva la asociación en la relación original (las diferencias entre porcentajes que aparecen en los espacios de propiedades de la relación bivariada decrecen notablemente al introducir una nueva variable). Esta última situación no implica que la relación original haya sido falsa, pero sí espúrea, puesto que en dicha covariación estaba incidiendo de algún modo la variable de control (Benson, 1974; Lazarsfeld, 1984; Mora y Araujo, 1984).

En el análisis de porcentajes de tablas multivariadas, el concepto de espureidad resulta central ya que pone de manifiesto una de las situaciones paradigmáticas en las que resulta útil la variable de control. A pesar del énfasis sobre la importancia de este concepto no debe perderse de vista que se trata más bien de una situación algo extrema que no suele ocurrir en la investigación social. Frecuentemente la variable de control no cambia sustantivamente la distribución, o bien hace que se atenúen las diferencias porcentuales de la relación original en una de sus categorías mientras que aumentan en otras.

Un modo de presentar estos contenidos puede consistir en esquematizar las situaciones posibles al introducir la variable de control (ver [Tabla 1](#)). Aunque en la práctica las variables de control habitualmente tengan el propósito de fortalecer una hipótesis que tiende a corroborarse –observándose una covariación–, también puede ser útil para el aprendizaje la comprensión de la posibilidad de controlar la ausencia de asociación, examinando si la misma persiste.

### *El desafío de pensar simultáneamente tres o más variables*

Es necesaria la lectura de cada relación parcial para analizar el efecto de la variable de control –que al igual que en las relaciones bivariadas se realiza comparando porcentajes en el sentido de la variable dependiente– y, a la vez, observar las similitudes entre las comparaciones realizadas en el cuadro de la relación bivariada original y los distintos cuadros generados con la introducción de la tercera variable. Uno de los pasos fundamentales en el aprendizaje de los estudiantes es la comprensión de que lo que en la lectura bivariada era la “conclusión” del análisis de una tabla, en la lectura multivariada pasa a ser el punto de partida de una nueva comparación en otro nivel. La introducción de dos variables de control supondría nuevamente un salto en complejidad y así sucesivamente. Una vez que se comprendió cabalmente cómo trabajar las tablas con una variable de control, puede ejercitarse la lectura de cuadros con más de una variable de control, sobre todo como una forma de explorar, tanto los docentes como los estudiantes, hasta qué nivel de complejidad se es capaz de llegar al pensar el inter-juego simultáneo de factores.

### *Las situaciones intermedias e indeterminadas*

El análisis del comportamiento de la variable de control muchas veces “decepciona” a los estudiantes que esperan modificaciones bien contundentes o, complementariamente, se encuentran ante la incertidumbre de no poder evaluar si las diferencias observadas entre las relaciones parciales son suficientes para hablar de un cambio o pueden tomarse como insignificantes. La comparación facilita pero no resuelve el problema de determinar cuánto hace relevante a una diferencia y cuando la misma puede considerarse despreciable. En esta situación, como en tantas otras, creemos que los docentes deben abstenerse de las “recetas” y enseñar a pensar en cada caso concreto cómo lo analizarían, teniendo en cuenta la teoría, el conocimiento acumulado sobre el tema, aspectos relativos a la cantidad total y la distribución de casos, entre otras cuestiones metodológicas siempre abiertas al debate y la reflexión en el aula. El desafío es desplazar el eje de las inquietudes desde la demanda de una respuesta cerrada a la reflexión siempre inacabada sobre las decisiones que deben tomarse y fundamentarse al emprender una investigación.

Por último, consideramos importante remarcar que todos los cuadros “dicen algo”, pues siempre hay una realidad a interpretar. Las lecturas que concluyen que una relación multivariada “no da nada” suelen ser inadecuadas, ya que aún no habiendo modificaciones en la relación original, la incorporación de la tercera variable contribuye en la construcción de conocimiento. Precisamente, que esta relación bivariada no se vea modificada por la incorporación de un nuevo efecto dice mucho en el sentido de fortalecer la relación original y comenzar a encontrar un camino hacia la corroboración de dicha hipótesis.

## **5. Breves conclusiones**

En la enseñanza del análisis cuantitativo recurrimos a distintas estrategias para superar dificultades en el proceso de aprendizaje y transmitir a los estudiantes que no hay fórmulas preestablecidas a la hora de realizar un análisis de datos (ver [Tabla 2](#)). Desde la elección de las variables a poner en relación y las decisiones

relativas a sus sistemas de categorías, hasta la interpretación de los datos numéricos, todo depende de los supuestos teóricos y del diseño propio de cada investigación. En la práctica, las hipótesis no se corroboran o refutan como en los esquemas didácticos del método científico sino que, en ciencias sociales, las asociaciones débiles, las corroboraciones parciales y los matices pueden ser lo más frecuente. Ello no significa que deba resignarse la búsqueda de conocimiento, sino que por el contrario se vuelve más necesaria la solvencia teórica para interpretar datos que no siempre son contundentes y que nunca hablan por sí mismos.

La capacidad de dominar las técnicas estadísticas y los procedimientos resulta imprescindible para realizar adecuadas interpretaciones, saber qué puede inferirse y cuáles son las limitaciones de un conjunto de datos. Aunque la construcción de las tablas de contingencia sea una operación que un programa de computación puede fácilmente realizar, las múltiples decisiones teórico-metodológicas son indelegables y requieren entrenamiento en el oficio. Aunque los recursos estadísticos permitan estandarizar las operaciones que se efectúan, no eximen de una disposición activa, flexible y creativa para el análisis. La mirada teórica y la imaginación sociológica no se dejan de lado al analizar datos cuantitativos sino que entran en juego en las decisiones procedimentales, en las observaciones y en las interpretaciones que se realizan, dotando de sentido, relevancia y coherencia a los datos.

## 6. Bibliografía

- Baranger, D. (1999). *Construcción y Análisis de Datos. Introducción al uso de Técnicas Cuantitativas en la Investigación Social*. Posadas: Editorial Universitaria de la Universidad Nacional de Misiones.
- Baranger, D. (2004). La construcción del objeto en La Distinción. En *Epistemología y metodología en la obra de Pierre Bourdieu*. Buenos Aires: Prometeo.
- Barton, A. (1973). El Concepto de espacio de propiedades en la investigación social. En Korn F. et al. *Conceptos y variables en la Investigación Social*. Buenos Aires: Nueva Visión.
- Benson, O. (1974). *El laboratorio de ciencia política*. Buenos Aires: Amorrortu.
- Blalock, H. (1986). *Estadística social*. México: Fondo de Cultura Económica.
- Bourdieu, P.; Passeron, J.C. y Chamboredon, J.C. (2011). *El oficio del sociólogo. Presupuestos epistemológicos*. Buenos Aires: Siglo XXI.
- Campbell, D. y Stanley, J. (1973). *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires: Amorrortu.
- Cohen, N. y Gómez Rojas, G. (2003). La lógica del experimento como instancia pedagógica. *Cinta de Moebio*, Facultad de Ciencias Sociales, Universidad de Chile, 18, pp. 219-226.
- Federico, A. y Oliva M. (1997). Participación femenina en el mercado de trabajo. El uso de loglineal en su estudio. En Salvia, A. (comp.). *Hacia una "estética plural" en la investigación social*. Buenos Aires: Oficina de publicaciones del CBC-UBA.
- García Ferrando, M. (1985). *Socioestadística*. Madrid: Alianza Universidad Textos.
- Hyman, H. (1984). El modelo del experimento y el control de las variables. En *El análisis de datos en la investigación social*. Buenos Aires: Nueva Visión.

Lago Martínez, S.; Mauro, M. y Álvarez, G. (2000). Análisis exploratorio multivariado. *Cinta de Moebio*, Facultad de Ciencias Sociales, Universidad de Chile, N.º 9, pp. 317-330.

Latiesa, M. (1991). Introducción a los modelos logarítmico lineales. *Papers Revista de Sociología*, N.º 37, pp. 97-112.

Lazarsfeld, P. (1984). El análisis de relaciones estadísticas. En Mora y Araujo, M. (comp.). *El Análisis de datos en la Investigación Social*. Buenos Aires: Nueva Visión.

López Roldán, P. y Fachelli, S. (2015). *Metodología de la Investigación Social Cuantitativa*. Barcelona: Dipòsit Digital de Documents Universitat Autònoma de Barcelona.

Mora y Araujo, M. (1984). El análisis de relaciones entre variables y la puesta a prueba de hipótesis sociológicas. En Mora y Araujo, M. (comp.). *El Análisis de datos en la Investigación Social*. Buenos Aires: Nueva Visión.

Sautu, R. (2005). *Todo es teoría. Objetivos y métodos de investigación*. Buenos Aires: Lumire.

Sautu, R. et al. (2010). *Manual de metodología. Construcción del marco teórico, formulación de los objetivos y elección de la metodología*. Buenos Aires: CLACSO.

Scribano, A. (2008). *El proceso de investigación social cualitativo*. Buenos Aires: Prometeo.

Zeisel, H. (1974). *Dígalo con números*. Buenos Aires: Fondo de Cultura Económica.

## Anexo de tablas

**TABLA 1.** Tipología de situaciones posibles al introducir la variable de control

	<b>Asociación bivariada</b>	<b>No asociación bivariada</b>
<b>La variable de control modifica la relación</b>	Espureidad - Especificación	Asociación encubierta
<b>La variable de control no modifica la relación</b>	Se fortalece la hipótesis original	No asociación que persiste

**TABLA 2.** Dificultades comunes en el aprendizaje de lectura de cuadros

<b>Dificultad</b>	<b>Tópicos y estrategias de enseñanza</b>
Comprensión del propósito del cuadro.	-Reflexión sobre las relaciones entre conceptos en la construcción de conocimiento. -Formulación de hipótesis.
Elaboración de hipótesis.	-Requisitos de las hipótesis. -Decisiones sobre los sistemas de categorías de las variables.
Contextualización de la lectura en la investigación, coherencia entre las hipótesis de cada cuadro y los objetivos e interrogantes.	-Elaboración del plan de análisis. -Reflexionar en cada caso cómo un cuadro responde a una necesidad de la investigación, en relación con la teoría y los interrogantes. El cuadro no es un fin en sí mismo ni en general suficiente por sí solo.



Comprensión de qué significan las frecuencias absolutas.	-Pasaje de la matriz de datos a la tabla de distribución conjunta de variables. -Importancia de considerar la cantidad de casos sobre la que se está predicando, en los totales y en las frecuencias condicionales.
Comprensión de para qué porcentualizar.	-Necesidad de estandarizar para hacer comparaciones válidas. -Comparación del uso de porcentajes calculados en distintos sentidos.
Aplicación de procedimientos de análisis comprendiendo el para qué, sin realizar lecturas mecánicas.	-Importancia de identificar distintas relevancias en los espacios de atributos según las hipótesis y la distribución de los datos. -Cuidar que la investigación no se subordine a la técnica. Las pruebas de hipótesis, los coeficientes y las lecturas porcentuales, así como todos los recursos estadísticos, son medios para la producción de conocimiento.
Lectura de los datos consciente de su construcción teórica, no como si fueran hechos dados.	-Reflexión sobre la construcción teórica de los datos, desde las posibles preguntas de cuestionario con las que se elaboraron las variables a las decisiones de codificación y procesamiento de la información. -Vigilancia epistemológica respecto a la evidencia construida.
Comprensión de qué implica que una relación no sea estadísticamente significativa.	-Diferencia entre significatividad estadística y relevancia teórica. -Factores que pueden influir en la significatividad estadística (decisiones de muestreo, construcción de variables, etc.). -Posibilidad de proseguir el análisis aunque no sea extrapolable al universo.
Selección del coeficiente.	-Formular distintas hipótesis asociadas a un mismo cuadro. -Analizar si cada una de ellas se corrobora al analizar la misma tabla de contingencia.
Interpretación adecuada del coeficiente.	-Romper el uso mecánico e individual del coeficiente. Analizar la relación entre la variable dependiente y múltiples independientes, observando cuál/cuáles contienen mayor asociación.
Comprensión de qué es controlar variables.	Lógica experimental y el análisis multivariado como alternativa "ex post facto".
Elección de variables de control.	Importancia de guiar teóricamente también las hipótesis alternativas.
Construcción del cuadro con variable de control.	Sentido de replicar la relación original bajo distintas condiciones, "como si" se tratara de un diseño experimental.
Lectura e interpretación del rol de la variable de control.	-Concebir a las situaciones de espureidad, especificación de la relación original y ajenidad de la variable de control como situaciones límite. -Reflexionar sobre los mecanismos que vinculan las variables y su secuencia temporal, confrontándolos con la nueva evidencia.
Análisis de los datos confiriéndoles sentido.	-Diferencia entre leer y analizar datos. -Identificar qué espacios de atributos son los más relevantes para cada hipótesis.